



500.41209X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): SEKI, et al  
Serial No.: 10 / 076,566  
Filed: FEBRUARY 19, 2002  
Title: INFORMATION RETRIEVAL DEVICE AND SERVICE

LETTER CLAIMING RIGHT OF PRIORITY

Assistant Commissioner for  
Patents  
Washington, D.C. 20231

MARCH 11, 2002

Sir:


Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s)  
the right of priority based on:

Japanese Patent Application No. 2001-091636  
Filed: MARCH 28, 2001

A certified copy of said Japanese Patent Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP

  
\_\_\_\_\_  
Carl I. Brundidge  
Registration No. 29,621

CIB/rp  
Attachment



W0138-01EF

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年 3月28日

出 願 番 号

Application Number:

特願2001-091636

[ST.10/C]:

[JP2001-091636]

出 願 人

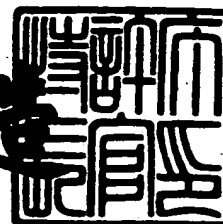
Applicant(s):

株式会社日立製作所

2002年 2月15日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

及 川 耕 造



CERTIFIED COPY OF  
PRIORITY DOCUMENT

出証番号 出証特2002-3007586

【書類名】 特許願

【整理番号】 K01005041A

【あて先】 特許庁長官殿

【国際特許分類】 G06F 15/00

【発明者】

    【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 ソフトウェア事業部内

    【氏名】 関 由美子

【発明者】

    【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 ソフトウェア事業部内

    【氏名】 斉藤 隆

【発明者】

    【住所又は居所】 東京都大田区大森北三丁目 2 番 1 6 号 株式会社日立システムアンドサービス内

    【氏名】 萩原 修身

【発明者】

    【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 ソフトウェア事業部内

    【氏名】 鬼頭 政義

【特許出願人】

    【識別番号】 000005108

    【氏名又は名称】 株式会社日立製作所

【代理人】

    【識別番号】 100075096

    【弁理士】

    【氏名又は名称】 作田 康夫

【手数料の表示】

    【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 検索システムおよび検索サービス

【特許請求の範囲】

【請求項 1】

計算機とデータベースとを接続したシステムにおける検索システムであって、  
ユーザからの検索要求を登録しておく検索予約登録部と、

前記登録内容にもとづいてデータベースに対して情報検索を行い、結果をユーザへ配信する検索システム部と、

前記データベースに対して情報検索を行う際に、重複した検索要求について、  
予め格納したルールに従って統合しユーザへ配信するデータを作成する検索処理部とを有することを特徴とする検索システム。

【請求項 2】

請求項 1 記載の検索システムであって、

前記検索予約登録において、現在および過去における検索要求の重複を調べて、  
検索条件の統合を行うことを特徴とする検索システム。

【請求項 3】

請求項 2 記載の検索システムであって、統合された検索条件に基づいて得られる検索結果を、各ユーザが所望する検索条件に基づいてユーザ別に結果を予め定める様式に従って編集して配信することを特徴とする検索システム。

【請求項 4】

請求項 3 記載の検索システムにおいて、ユーザ別に結果を編集する際、ユーザが所望した検索結果内容に重複がある場合には重複を排除し、統合して編集した結果を配信することを特徴とする検索システム。

【請求項 5】

請求項 2 または請求項 3 記載の検索システムであって、前記検索予約登録において過去の検索結果を少なくとも一つ以上保持しておき、ユーザからの検索要求を行う際に、保持しておいた検索結果の中から検索結果を求めるか、または新規に原データからの検索を行うかをユーザが選択する方法を具備したことを特徴とする検索システム。

【請求項 6】

請求項 3 記載の検索システムにおいて、ユーザ間で重複する検索要求があって、それらの要求を統合して検索結果を得た場合、得られた検索結果を要求した各ユーザに複製して配布することを特徴とする検索システム。

【請求項 7】

請求項 1 または請求項 2 または請求項 3 または請求項 4 または請求項 5 または請求項 6 記載の検索システムにおいて、前記検索予約登録にもとづく検索を予め定められた時間間隔で取得または配信することを特徴とする検索システム。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明はネットワーク上のデータベースに蓄積された情報の検索方法及び装置に関する。

【 0 0 0 2 】

【従来の技術】

(1) 蓄積された文書に対する予約検索の従来技術としては特開平 7 - 3 3 4 5 2 2 号公報にあるように、予め登録された検索条件と検索時機に基づいて該当する文書の一覧表を取得する技術がある。

【 0 0 0 3 】

(2) また、複数のユーザからなされた複数の検索要求を一括して処理する方法は特開平 6 - 6 0 1 2 1 号公報に示されている。

【 0 0 0 4 】

【発明が解決しようとする課題】

上記従来の技術 (1) においては、検索結果は文書一覧として与えられるので、検索内容をメールにて通知し、また通知内容に重複があった場合の処理については触れていない。

【 0 0 0 5 】

また、上記従来の技術 (2) においては、複数のユーザがそれぞれ検索要求を行った際にこれらを統合して検索処理を行い、検索効率を上げる方法と、その結

果を各ユーザに展開する方法について述べられているが、不特定多数のネットワーク上のデータベースに対して検索処理を行う場合の効率化の手法について、また結果を各ユーザへ展開する際の具体的な実現方法については述べられていない。

#### 【0006】

##### 【課題を解決するための手段】

本発明では、計算機とデータベースとを接続したシステムにおける検索システムであって、ユーザからの検索要求を登録しておく検索予約登録部と、前述の登録内容にもとづいてデータベースに対して情報検索を行い、結果をユーザへ配信する検索システム部と、前述のデータベースに対して情報検索を行う際に、重複した検索要求について、予め格納したルールに従って統合しユーザへ配信するデータを作成する検索処理部とを有する。

#### 【0007】

尚、本発明ではネットワーク上にある不特定多数のデータベース上に蓄積された文書に対して、予め定められた検索条件、検索時機において予約検索を行う手段と、少なくとも一人以上のユーザが個々の検索条件登録を行い、その登録を保持して活用することができる手段と、ユーザ間で検索条件の登録内容に重複があった場合には重複を統合して検索条件の作成を行い検索する手段と、得られた検索結果について、ユーザ間の重複分について各ユーザが所望する検索条件登録に合致するように情報の複製を行って各ユーザの検索条件登録ごとに展開する手段と、展開した内容中に重複があった場合はこれを排除してユーザが所望する形態に編集し、編集結果をユーザが予め登録したメールアドレスあるいは配信先へ配信する手段とを備えた情報検索システム、もしくはこのような検索システムを利用した検索サービスからなる。

#### 【0008】

##### 【発明の実施の形態】

以下、本発明の実施例について図面を用いて説明する。

#### 【0009】

図1は本発明を適用した検索システムの構成と検索サービスの動作を説明する

図である。本システムでは複数の検索対象データベースシステム（DBMS）1、2、、、、nがネットワークを介して接続され検索が可能である環境において、本システムのユーザ1、2、3、、、nが予め登録しておいた検索条件に従って検索された結果をメールなどの媒体を通じて配信されるサービスを受けることができる。

#### 【0010】

以下に、本発明による検索結果配信サービスを提供するプログラム・各システムの構成について説明する。

#### 【0011】

103および104の検索予約登録部（A）（B）はそれぞれユーザが予め設定するユーザ別配信要求102によって予約検索条件を登録しておくものである。これはユーザの数だけあってもよいが、後で述べるように、ユーザ1がユーザ2の予約検索条件をそのまま再利用してもよい。その場合両ユーザの検索予約登録部は別々にあってもよいが、資源節約のために同じものとなってもよい。

#### 【0012】

検索予約登録部に登録された内容で予約検索を行うのが情報検索部105である。情報検索部105は次の構成からなる。106検索条件統合部は各検索予約登録部の検索条件の重複を調べて重複があればこれを統合して検索式作成部107に結果を渡す。

#### 【0013】

検索式作成部107は与えられた検索条件に基づいて検索式を作成する。作成された検索式は108検索実行部においてコマンド発行される。109前検索結果記憶参照部において発行されたコマンドが既に検索済みの結果を参照する場合はこれを参照し、そうでない場合は新規に検索を行う。110検索結果取得部に検索結果を取得し、111検索結果配信部において各検索予約登録部結果データベース112および113への検索結果の配信を行う。

#### 【0014】

114ユーザ別配信振り分け処理部では各ユーザに配信するための検索結果の振り分けを行い、同じユーザに配信する結果内に重複する結果が含まれていれば



これを統合して配信する（115）。

【0015】

図2でユーザ別配信要求102によって検索予約登録を行う方法についてユーザインターフェースの画面例を用いて説明する。

【0016】

201はユーザAが利用する検索予約登録画面の例である。ここでは基本的な機能についてのみ説明する。202は検索カテゴリを選択する項目である。例えば、ユーザの使用頻度の高い検索条件については予めこのようにカテゴリ項目を列挙してユーザに選択させることで、ユーザの負担を軽減することができる。この例では「インターネット」と「テキスト検索」が選択され検索予約されている。その他の検索条件については203に示すようにユーザが検索したい言葉を入力する欄を設け、単語や文章による指定を受け付ける。

【0017】

また、この欄では、うろ覚えの言葉や、綴り字の間違った場合に文章校正を行う手段を補助的に設けておいてもよい。202や203でユーザが自分の検索したいキーワードをうまく設定できないとき、あるいは、既に誰か別のユーザが設定した検索条件を再利用したいときには、204で別ユーザの検索予約登録情報を参照することができる。

【0018】

これは、予めユーザAが参照することが許されているユーザ群について、それらのユーザが設定した検索条件を選択し、参照すると、参照されるユーザが設定しているカテゴリやキーワードが表示され各項目が再利用可能となる。

【0019】

ユーザAはそれをそのまま利用して登録してもよいし、さらにその情報に新たな項目を追加・削除して編集してもよい。205で検索結果の配信先を指定する。この例ではデフォルト設定では社内のメールアドレス宛てとなっているが、個人用のメールアドレスに配送したり、FAXへ転送するなどの設定をしてもよい。206は情報の鮮度オプションの指定である。ここで情報の鮮度とは、検索対象となるデータの新しさを示している。

## 【0020】

例えば、予約検索の場合、1週間ごと、あるいは1日ごとの定められた時機に検索を行い、この結果を全文検索用データベースに登録しておいて、この登録済みデータに対してアクセスすることによって処理の高速化を計っている場合がある。そこで、情報の鮮度オプションを設け、指定がなかった場合（デフォルト）はこの予め登録してあるデータに対して検索を行い、結果を取得することとする。しかし場合によっては最新のデータを新規に検索したいこともある。

## 【0021】

このような場合のために情報の鮮度オプションを設け、ユーザがどの程度まで新しい情報を必要としているかを選択することができる。例では、24時間以内に作成済みの検索用データベースに対して検索を行う場合と、新規に原データ（ネットワーク上の検索対象データベースシステム群）に対する検索を別途行う場合の2種類の選択肢を表示している。以上202～206の情報を指定して検索予約登録を行い、その情報は検索予約登録部（103）に格納され管理される。

## 【0022】

図3に106検索条件統合部の処理フローを示す。各検索予約登録部（103，104）に登録されている検索条件の記述内容に重複がないか調べる（301）。重複がある場合は重複するキーワードから重複リスト（401）を作成する（302）。重複を排除して統合した検索条件を検索要求として検索式作成部107へ渡す（303）。また処理301で重複がなかった場合はそのまま処理303へ進む。

## 【0023】

図4に重複リスト（401）の例を示す。この例においては、キーワード「インターネット」で予約検索登録しているユーザはユーザA、ユーザB、ユーザCの3人で、それぞれの登録している検索予約登録部はユーザAが検索登録部Aであり、ユーザBとユーザCが検索登録部Bを共有している。また、キーワード「テキスト検索」で予約検索登録しているユーザはユーザAとユーザBの二人である。

## 【0024】

図4の重複リストにある重複キーについては検索条件統合部106において統合される。すなわち、「インターネット」に対して検索を要求しているユーザは三人いるが、検索そのものは一度でよい。また「テキスト検索」についても要求ユーザが二人いるが、検索そのものは一度でよい。そこでこれらを統合して、「インターネット」と「テキスト検索」の一回の検索として検索式を作成するのが107の検索式作成部である。作成された検索式は108検索実行部でコマンドとして発行される。

## 【0025】

ここで発行されたコマンドはそのまま実行してもよいが、本実施例では引き続き前検索結果記憶参照部109の処理を行う。このように検索の重複を統合することで検索用サーバの処理を高速化し、負荷を軽減することができる。106検索条件統合部の統合方法、107検索式作成部の作成方法、108検索実行部の実行方法について詳細は特開平6-60121号公報にあるような方法を利用してよい。

## 【0026】

図5に前検索結果記憶参照部109の処理フローを示す。検索実行部108から出された検索要求の鮮度オプション指定が新規検索を指定しているかどうかを判定し(501)、YESであればネットワーク上の検索対象データベースシステムに対して検索を実行し、検索結果を最新の前検索結果として保存する(503)。処理501でNOの場合、すなわち新規検索指定がなされていなかった場合、すでに格納されている前検索結果で、指定された鮮度オプション(24時間以内の情報、1週間以内の情報など)に該当するデータに対して検索を実行する(502)。

## 【0027】

なお、前検索結果記憶参照部109が前検索結果を対象として検索実行する手法については、特開平6-60121号公報にあるように複数の検索対象データベースがあった場合にどのデータベースを対象に検索処理を行うか指定して選択的に検索処理を行う方法を利用してよい。

## 【0028】

110 検索結果取得部で検索された結果を取得し、111 検索結果配信部で各検索予約登録部結果データベースへ結果を配信する。

【0029】

図6に111 検索結果配信部の処理フローを示す。

【0030】

図6に示す検索結果配信部111では、まず重複リスト401を参照して、検索結果のキーと重複キーとを比較し、重複リストに載っている検索予約登録分だけを必要な数だけ複製する(601)。複製された検索結果を各検索予約登録部結果データベース(112、113)へ配布して登録する(602)。

【0031】

図7にユーザ別配信振り分け処理部114の処理フローについて示す。各ユーザが予め設定した検索予約登録を参照して、検索予約登録結果データベースからユーザ所望の検索結果を取得して内容を記憶する(701)。この内容をもとに検索結果重複統合処理(702)にうつる。

【0032】

図8に検索結果重複統合処理(702)の処理フローを示す。まず検索結果内容に重複する結果が含まれていないか判定する(801)。判定に際しては重複度判定ルール群802を判定基準として利用する。

【0033】

重複度判定ルール群とは、例えば、一致度定義ルール803、類似度定義ルール804、同一情報起源定義ルール805、その他定義ルール806などからなる。具体的には、一致度定義ルールとは、検索結果がそれぞれ一致するもの、あるいは、互いに全文検索結果の一致度が80%を超えるものは一致すると見なす、といったような定義がなされている。

【0034】

また、類似度定義ルール804とは、文書の内容を要約したときにそれぞれの検索結果の要約が80%以上類似している場合、それらの結果は類似文書であって、重複と見なす、といったような定義がなされている。また、同一情報起源定義ルールとは、検索結果を取得したデータの起源が同じところから発生している

もの（例えば国外にあるA社の社外発表をB新聞社が翻訳して紹介した記事と、C新聞社が翻訳して紹介した記事とは本質的に同じ内容を示す、など）は重複と見なす、といったような定義がなされている。

【0035】

その他定義ルール806とは本検索システム側で個別に定義することのできるルールを示す。以上のような重複度判定ルール群802を基にして、重複する結果がないかどうか判定した結果、重複する結果が含まれているときには、結果優先度ルールに基づいて結果を統合する（807）。

【0036】

結果優先度ルール808とは、重複する結果が含まれていたときにどちらを優先してユーザに配信し、どちらを重複データとして削除するかを決めるためのルールである。先ほどの例でいえば、B新聞社の翻訳技術とC新聞社の翻訳技術ではB新聞社の方が的確であることが予めわかっているとき、B新聞社の記事を優先して、C新聞社の記事を結果から重複として削除する、などを定めるものである。

【0037】

また、例えば、「インターネット」および「テキスト検索」で検索した結果と、「インターネット」のみで検索した結果、「テキスト検索」のみで検索した結果はそれぞれ包含関係にあり重複している。

【0038】

従って、検索結果は「インターネット」と「テキスト検索」の両方を含むものを優先し、その結果、両方を含む検索結果として選択された情報と同じ情報については、「インターネット」のみ、および「テキスト検索」のみで検索した結果からは、すでに結果が得られた重複データとして削除する、などを定めるものである。

【0039】

702 検索結果重複統合処理で重複分を削除して残った検索結果について、配信テンプレート（1101）を参照して検索結果を編集して整形し、ユーザが予め指定したアドレスへ結果を配信する（703）。

## 【 0 0 4 0 】

図 9 に以上で述べてきた検索結果データ重複統合の仕組みをまとめて示す。検索対象データ群（検索対象 DBMS 1、検索対象 DBMS 2）はそれぞれネットワーク上に接続された不特定多数の原データであっても良いが、検索対象として予め検索して登録しておいた前検索結果データベースであっても良い。この検索対象群に対して検索処理を行ってデータ 1 およびデータ 2 が得られたものとする。

## 【 0 0 4 1 】

この時、検索処理は重複を排除して統合されているので、検索対象データ群に対するアクセスは最小限に抑えられ、検索効率や検索用サーバの負荷の低減がなされている。また、検索対象データが、コンテンツ利用に関して課金制度を用いていた場合、例えば検索対象 DBMS 1 で 1 度の検索にかかる費用 XX 円であれば通常 10 回検索を行うと XX 円の 10 倍の課金が行われるが、このシステムでは検索回数が最低限に抑えられているので課金も少ない。

## 【 0 0 4 2 】

ただし、不正な使用をさけるために、検索対象 DBMS 側と本検索サービス側とで別途契約を結んで運用することが望ましい。例えば後で複製する数に関する記述部分（重複リスト 4 0 1 など）のみを互いに共有しておいて、いくつ複製が作られたかによって自動的に課金を増やすことで、課金制度に支障をきたすことなくお互いのシステム負荷を低減する方法などが可能である。

## 【 0 0 4 3 】

また、ユーザ側で予算を定めておき、課金と予算の関係から検索対象 DBMS 1 からデータ取得するか、検索対象 DBMS 2 からデータ取得のかなどを選別する方法も可能である。得られた検索結果データ 1 とデータ 2 とは検索結果配信部 1 1 1 によって必要な数だけ複製が作られ、検索予約登録部（A）結果 DB に登録される。

## 【 0 0 4 4 】

この登録内容をユーザ別に配信振分け処理を行う中で、検索結果重複統合処理 7 0 2 が行われてデータ 1 とデータ 2 の中に存在する内容重複部分 9 0 5 を重複

排除して統合し、統合された配布用データをユーザ別結果配信 115としてユーザに配信する。

## 【0045】

図10に具体的な重複排除の例を示しておく。

ユーザAの予約検索結果がデータ1とデータ2であった場合、この図の例ではデータ1中に「インターネット」と「テキスト検索」の両方を含む検索結果(1) 2月10日のA新聞社記事が得られている。

## 【0046】

一方、データ2中に「テキスト検索」を含む検索結果として同じ(1) 2月10日のA新聞社記事が得られている。それぞれは内容が同じで重複しているので、このままユーザに配信してしまうと配信内容が冗長になり、ユーザとして読みづらく、わずらわしいものになってしまう。

## 【0047】

そこで、内容が重複している部分905を検索結果重複統合処理702によって統合した配布用データを作成している。この図の例ではデータ1の検索結果のみを採用して、データ2に含まれている同じ内容の検索結果は排除されている。この統合された配布用データをユーザ別結果配信として配信する。

## 【0048】

図11に配信用テンプレートの例1101を示す。この例では、検索に使用したキー、検索日時、目次、検索結果、ヒット数、ヒットした数だけの検索結果の情報(タイトル、作成者、作成日付、URL、概要、コメントなど)を繰り返し、最後に終了コメントを出力する書式としている。なおここでいう検索日時とは、新規検索の場合は検索日時そのものであるが、予め登録されていた前検索結果を参照した場合はその前検索結果が検索された日時をユーザに明らかにするものである。

## 【0049】

図11のテンプレートの中のコメントは、検索結果や類似度定義ルールに基づいた類似度の比較結果に基づいて作成しても良い。

## 【0050】

たとえば、雑誌社αのインターネットに関する記事と雑誌社βのインターネットに関する記事が95%類似という類似比較の結果がでた場合には「雑誌社αのインターネットに関する記事と雑誌社βのインターネットに関する記事の類似度は95%でした。本結果通知には雑誌社αの記事を載せ、雑誌社βの記事を省略しました。」というように、どの記事を比較したか、どの記事を省略したかがわかるようなコメントをつけてもよい。その際に、文章のテンプレートを予め格納しておき、類似度の数値や検索ヒット件数、情報の出所（新聞社、URL、学会名等の情報）、記事や論文のタイトル等については類似判定結果や検索結果等の値を入れるようにしてもよい。

## 【0051】

また、コメントは検索式ごとの検索結果に応じてつけてもよいし、コメントを希望する結果のみにつけてもよいし、検索結果送信ごとにつけてもよい。

## 【0052】

なお、記事の省略に関して、予めどの記事を結果に載せ、どの記事を省略するかを重複度判定ルール群802の中で定義しておいてもよい。たとえば、記事がA新聞社の記事とB新聞社の記事とが類似度85%である場合には、「A新聞社の記事を載せ、B新聞社の記事を省略する。」という定義をしてもよい。この場合は予め定義されたルールに従ってコメントを自動生成するようにしてもよい。

## 【0053】

さらに、複数の記事の内容の重複が複数あったなどには、記事を載せる優先順位を定めておいても良い。たとえば、「A新聞社、B新聞社、C新聞社の3社の記事で重複があった場合には、A新聞社の記事を載せ、B新聞社およびC新聞社の記事は省略。B新聞社、C新聞社の2社の記事があった場合はB新聞社の記事を載せ、C新聞社の記事は省略。」という定義を定めることで、どの記事を載せ、どの記事を省略するかを重複度判定ルールで定めておき、この内容に従ってコメントを記載してもよい。

## 【0054】

なお、これらはあくまでも一例であり、別途コメントのテンプレートを用意しておき、コメントのテンプレートと重複度判定ルール群802とを利用した文章



を記載しても良いし、他の方法でもよい。

【 0 0 5 5 】

図 1 2 に 1 1 0 1 のテンプレートを用いて配信されるメールの例を示す。ユーザ A が受け取るメール 1 2 0 1 は、まずユーザ A が指定したキーとして「インターネット」「テキスト検索」、次に検索日時は 2 0 0 1 年 3 月 1 日のデータを使用し、検索結果として、まず「インターネット」と「テキスト検索」の両方を含むもののヒット数が 2 件、以下その 2 件の情報、次に「テキスト検索」を含むものの検索結果ヒット数 1 件の情報が記述されている。

【 0 0 5 6 】

なおここではすでに「インターネット」と「テキスト検索」の検索でヒットした ( 1 ) と ( 2 ) の 2 件の情報については重複するので削除されている。さらに「インターネット」を含むものの検索結果ヒット数 1 5 2 3 件と続いている。この例にあるように、ヒット数が多すぎてメールの受信容量を越える場合にはコメントを記述して概要の出力を控える。ここでは「メール容量オーバーのため、検索結果 DB 格納エリア 1 を直接参照してください」としてメールが冗長になることを防いでいる。

【 0 0 5 7 】

このような結果が多すぎて容量オーバーなどが生じる場合にはメールの最後尾にこの情報を記載することでユーザ側の可読性を上げ、情報取得を素早く行えるようにするなどの工夫の余地がある。

【 0 0 5 8 】

以上説明したように本実施例においては、予めユーザが登録しておいた検索予約登録に基づいて予約検索を行うにあたり、ユーザ間で重複する検索条件については統合して検索式を作成して検索を行うので、検索用サーバおよびネットワーク上の不特定多数の検索対象データベースシステムおよびそれらが接続されるネットワークにおける処理の高速化、負荷の軽減を図るという効果を得ることができる。

【 0 0 5 9 】

また本実施例によれば情報の鮮度オプションを設定して予約検索を行うことが

できるので、予め登録されているデータに対して検索を行う場合と、新規にデータの検索を行う場合を区別して処理でき、検索用サーバおよびネットワークの処理の高速化と負荷の軽減を図るという効果が得られる。

#### 【0060】

なお、検索対象データベースがデータ取得について課金制度をとっていた場合には、特に制約がなければ課金を減らすという経済効果が得られるが、データ使用に制約がある場合には、重複リストに基づいて、複製して使用するデータ量だけの使用料を支払う、あるいは別途データ利用契約を結ぶとすればよい。

#### 【0061】

また、一致度定義ルール803、類似度定義ルール804、同一情報起源定義ルール805、等を用いることで、異なるDBからの検索結果から、検索ヒット率の高いファイルを抽出し、前述の抽出したファイルの内容を要約類似度で比較し、前述の比較結果に応じて結果を編集し、編集した結果をテンプレート1101に基づいてフォーマットの変換をすることや、検索結果や類似度の比較結果に応じたメッセージをつけることが可能となり、検索結果をまとめる労力を軽減することが可能となる。

#### 【0062】

##### 【発明の効果】

以上説明したように、本発明によれば、ネットワーク上にある不特定多数のデータベース上に蓄積された文書に対して予約検索を行うに際し、少なくとも一人以上のユーザユーザ間で検索条件の登録内容に重複があった場合には重複を統合して検索することで、検索処理の効率を上げることができるという効果が得られる。

#### 【0063】

また、得られた検索結果について、ユーザ間の重複分について各ユーザが所望する検索条件登録に合致するように情報の複製を行い、展開した内容中に重複があった場合はこれを排除してユーザが所望する形態に編集し、編集結果をユーザが予め登録した配信先へ配信することができるという効果が得られる。

##### 【図面の簡単な説明】

【図 1】

本発明のシステム構成とシステム全体の動作を示した図である。

【図 2】

本発明の一実施例における検索予約登録の画面例を示した図である。

【図 3】

本発明の一実施例における検索条件統合部の処理フローを示す図である。

【図 4】

本発明の一実施例における重複リストの例を示す図である。

【図 5】

本発明の一実施例における前検索結果記憶参照部の処理フローを示す図である。

【図 6】

本発明の一実施例における検索結果配信部の処理フローを示す図である。

【図 7】

本発明の一実施例におけるユーザ別配信振り分け処理部の処理フローを示す図である。

【図 8】

本発明の一実施例における検索結果重複統合処理のフローを示す図である。

【図 9】

本発明の一実施例における検索結果データ重複統合の仕組みを説明するための図である。

【図 10】

本発明の一実施例における重複統合された検索結果データの内容の例を示す図である。

【図 11】

本発明の一実施例における配信用テンプレートの例を示す図である。

【図 12】

本発明の一実施例における配信用テンプレートを用いた配信メールの例を示す図である。

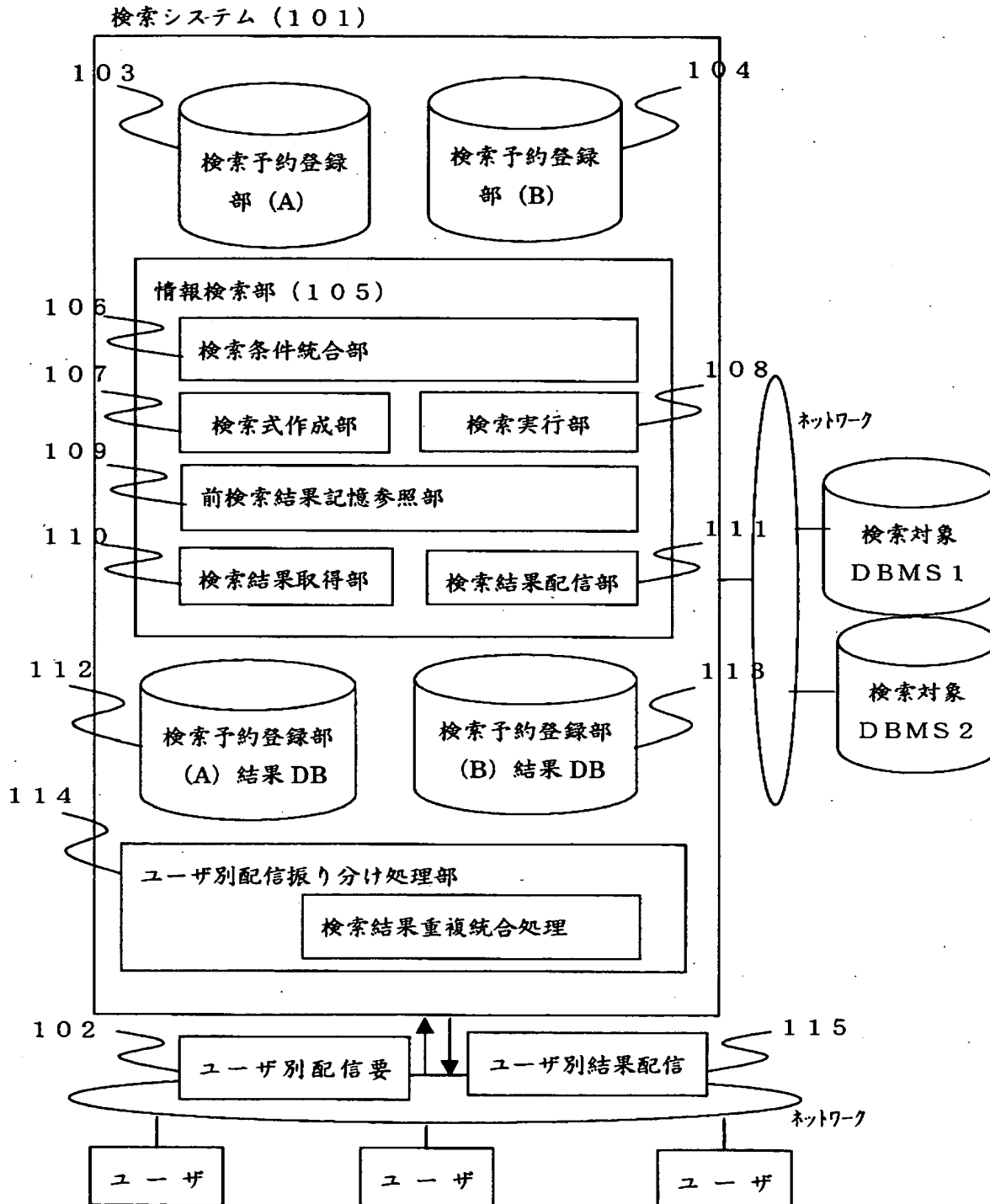
【符号の説明】

1 0 3 ・ 1 0 4 検索予約登録部、 1 0 6 検索条件統合部、  
1 0 7 検索式作成部、 1 0 8 検索実行部、  
1 0 9 前検索結果記憶参照部、 1 1 0 検索結果取得部、  
1 1 1 検索結果配信部、  
1 1 2 ・ 1 1 3 検索予約登録部結果データベース、  
1 1 4 ユーザ別配信振り分け処理部、  
2 0 1 検索予約登録画面、 4 0 1 重複リスト、  
8 0 2 重複判定ルール群、 8 0 3 一致度定義ルール、  
8 0 4 類似度定義ルール、 8 0 5 同一情報起源定義ルール、  
8 0 8 結果優先度ルール、  
1 1 0 1 配信用テンプレート、  
1 2 0 1 配信メール。

【書類名】 図面

【図 1】

图 1



【図2】

図2 検索予約登録画面の例

201

202 ユーザAさんの予約検索登録画面

○ご希望のカテゴリを選んでください

☒ インターネット  
☐ データベースシステム  
☒ テキスト検索  
☐ ハードウェア

203 ○ご希望のキーワードを入力してください

204

○グループ内の別ユーザを参照しますか？

205 参照するユーザ  ☒

○配布先を指定してください（デフォルト社内アドレス）

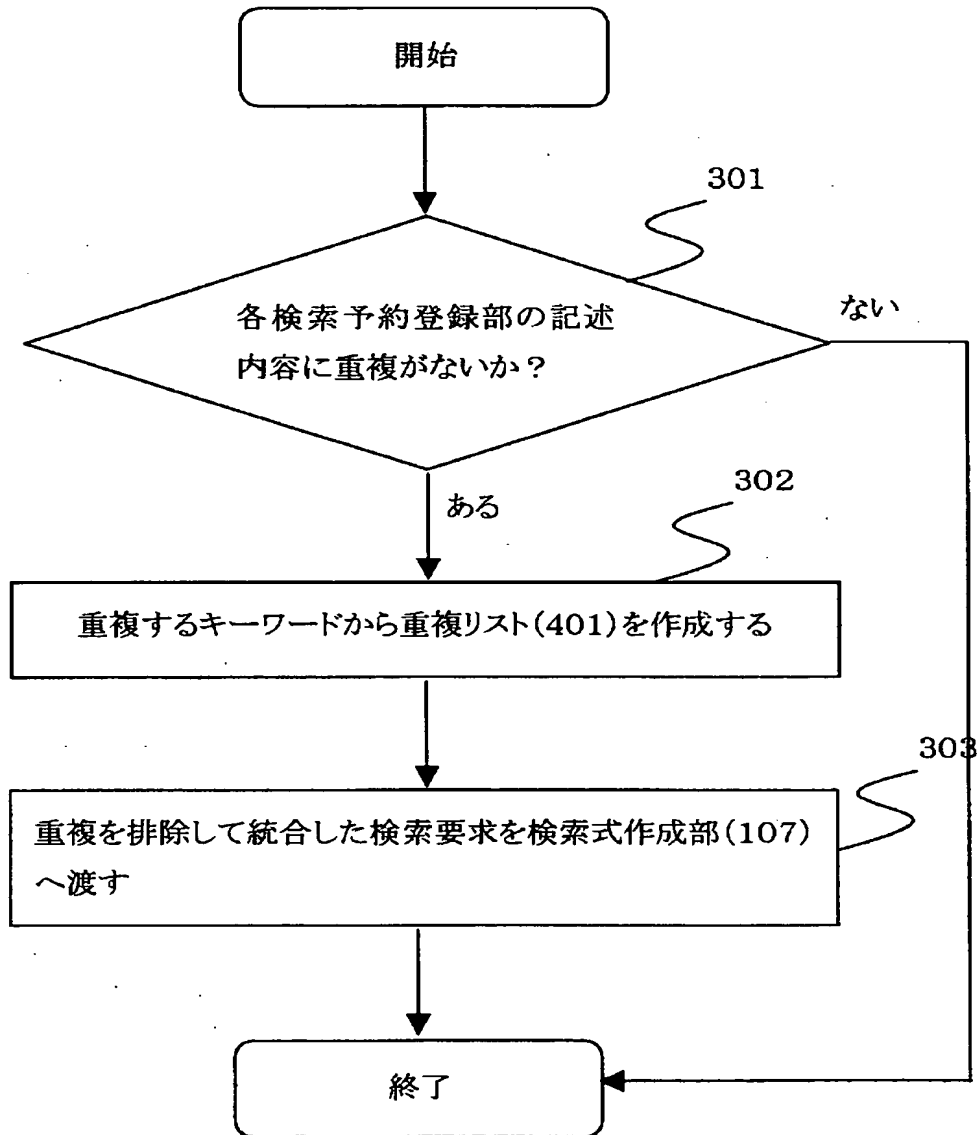
206

○情報の鮮度オプション指定（デフォルトは1週間以内）

☐ 24時間以内 ☐ 新規検索

【図3】

図3 検索条件統合部(106)



【図 4】

図4 重複リスト(401)の例

重複キー	要求元検索予約登録	要求ユーザ
インターネット	検索予約登録部 A	ユーザ A
インターネット	検索予約登録部 B	ユーザ B
インターネット	検索予約登録部 B	ユーザ C

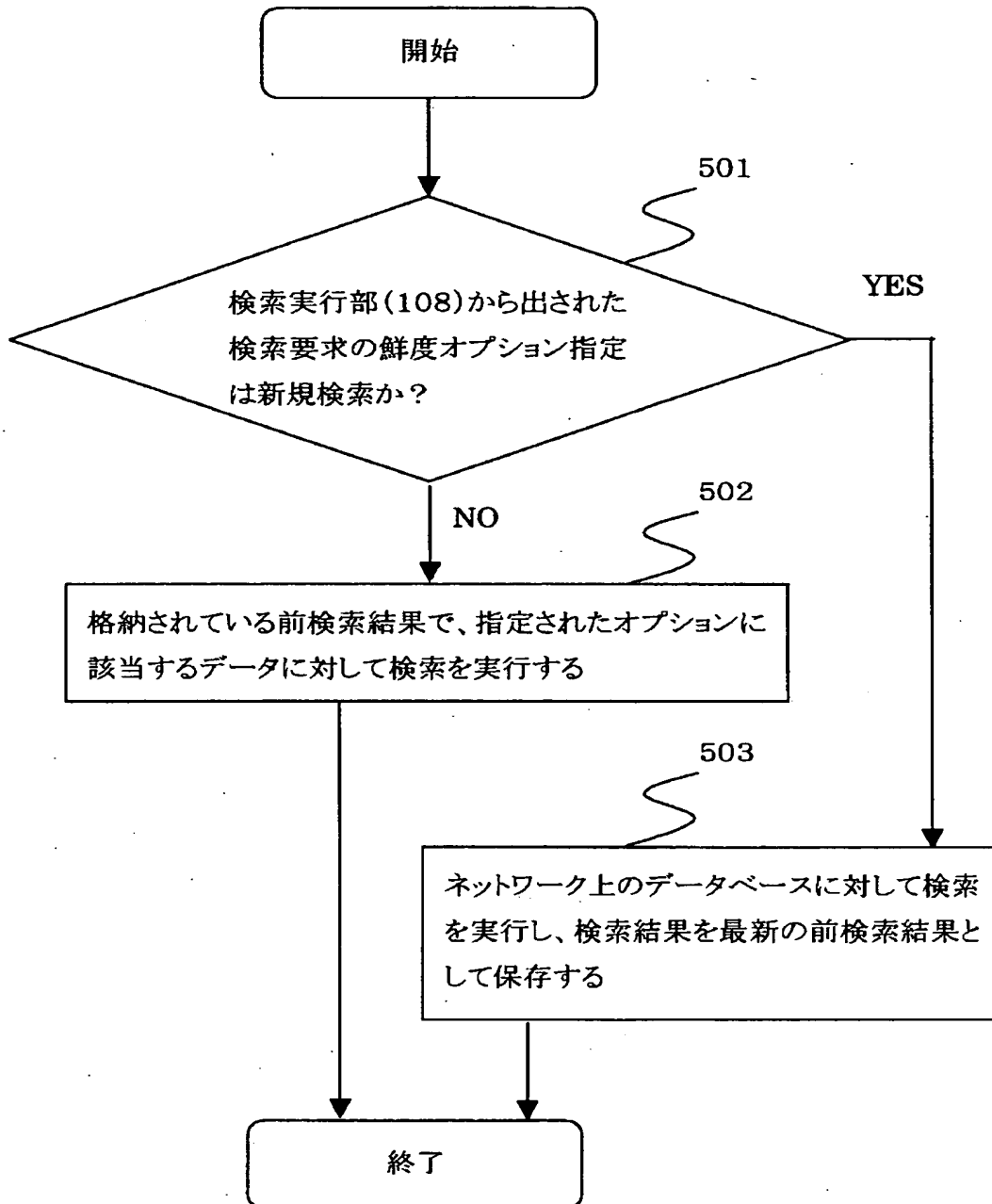
  

重複キー	要求元検索予約登録	要求ユーザ
テキスト検索	検索予約登録部 A	ユーザ A
テキスト検索	検索予約登録部 B	ユーザ B



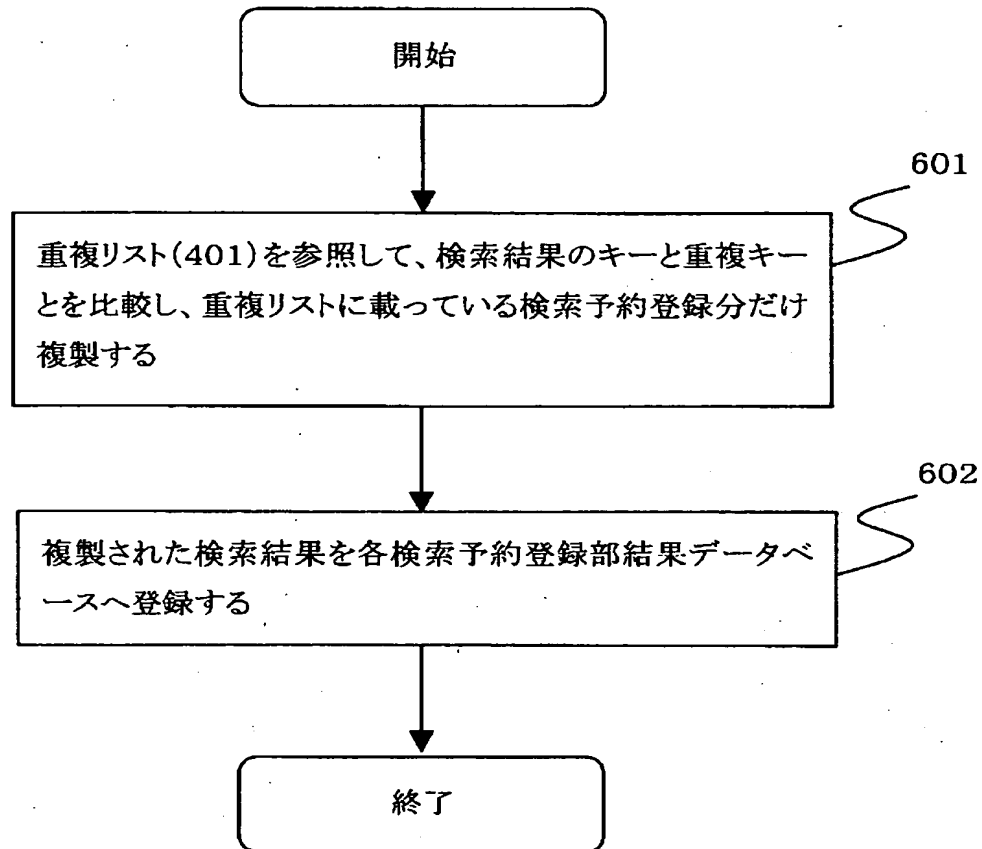
【図5】

図5 前検索結果記憶参照部(109)



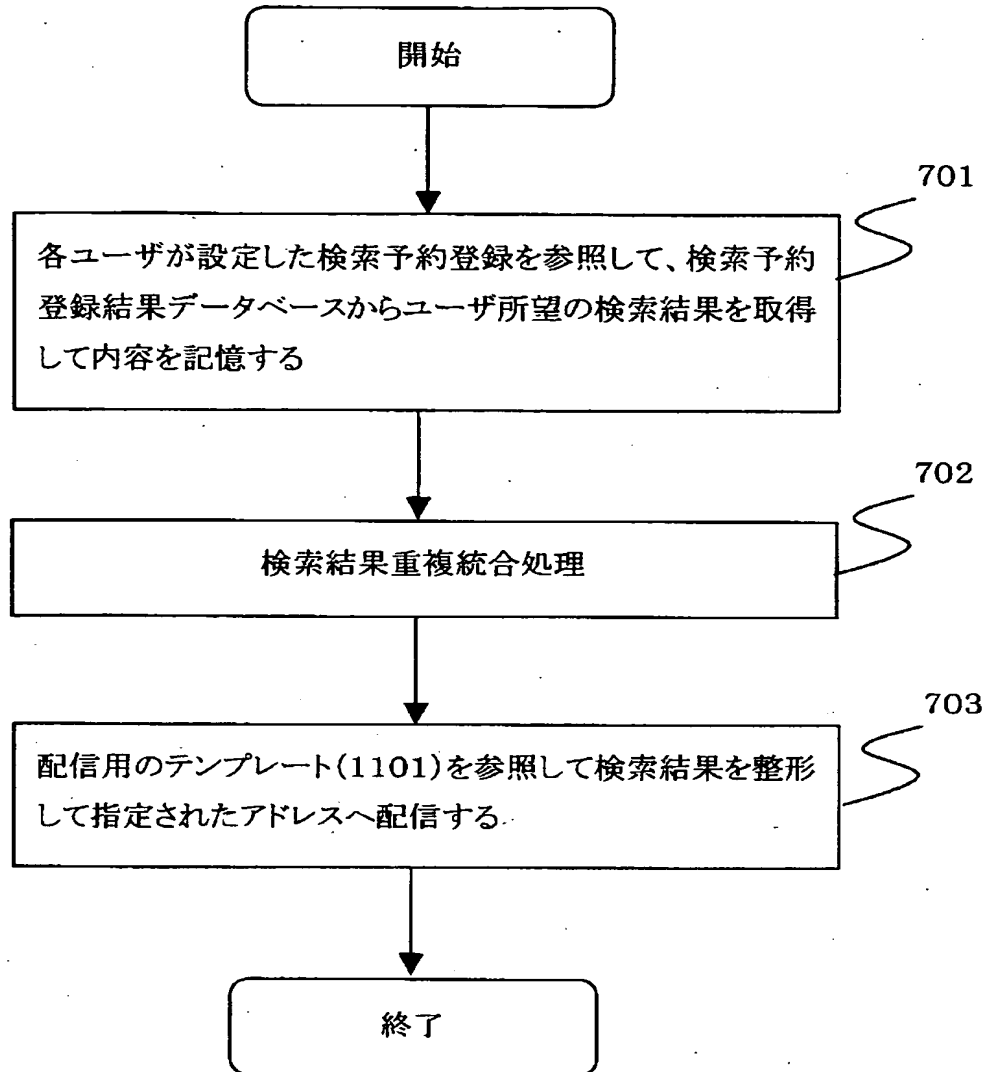
【図6】

図6 検索結果配信部(111)



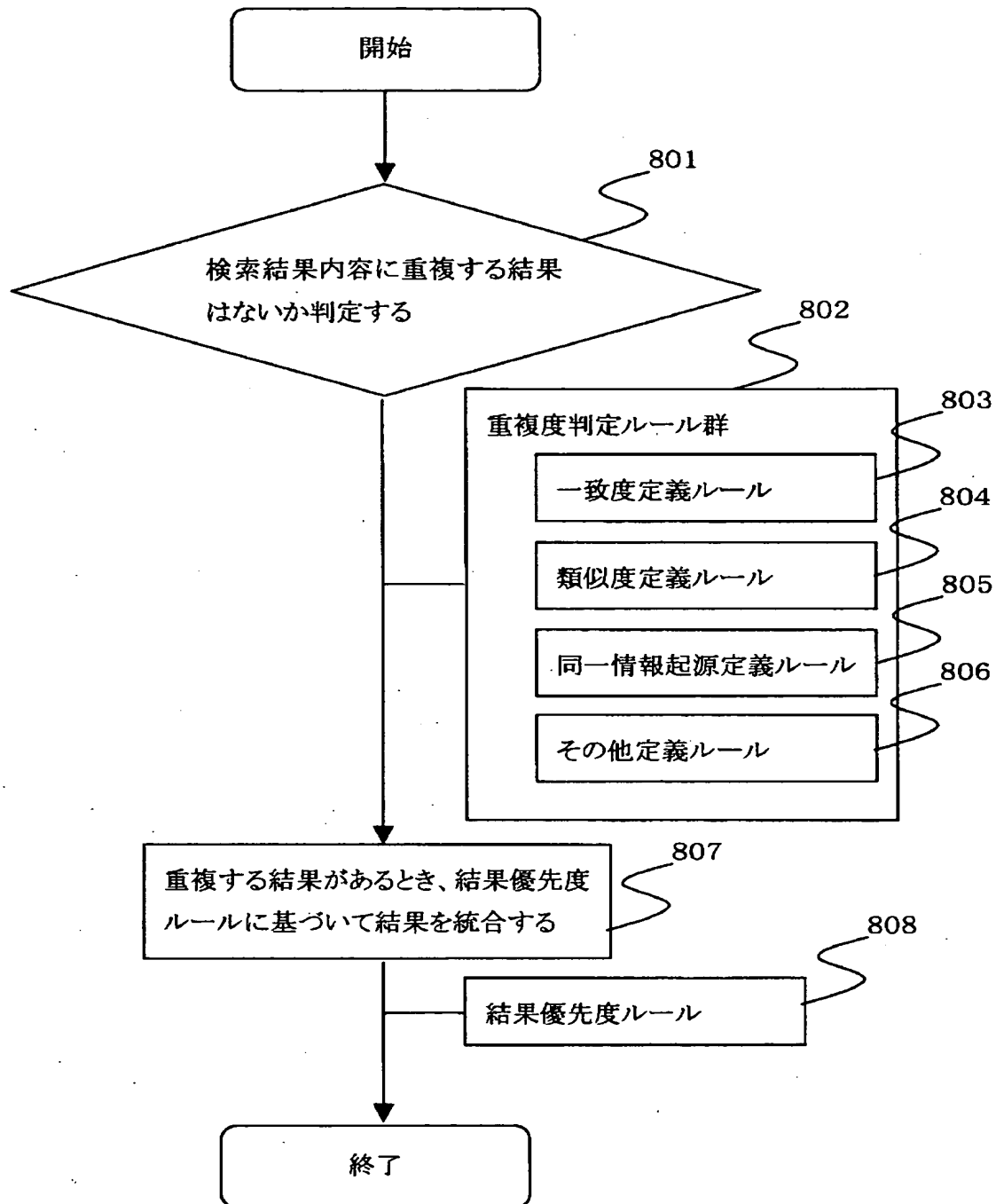
【図7】

図7 ユーザ別配信振り分け処理部(114)



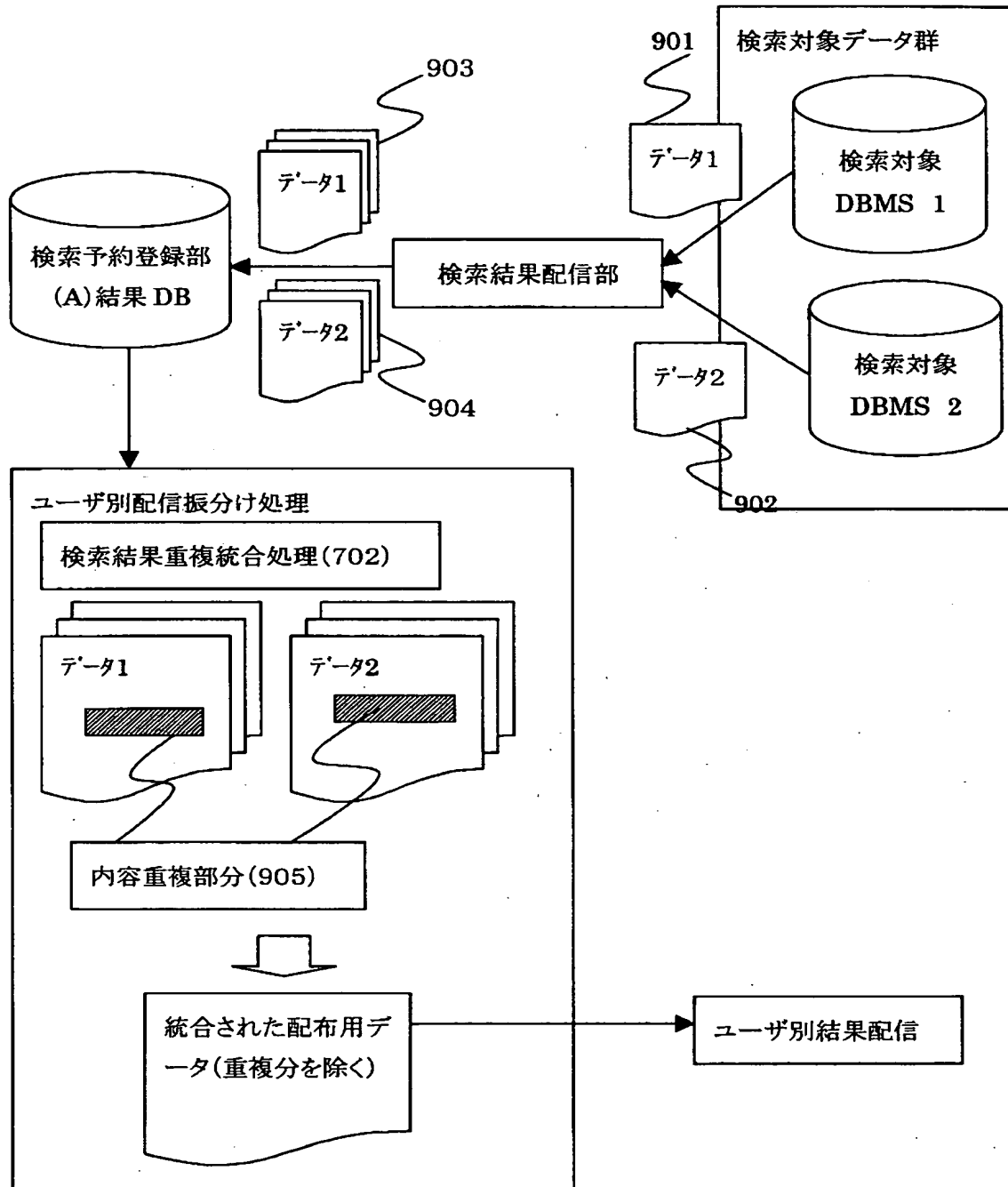
【図 8】

図 8 検索結果重複統合処理(702)のフロー



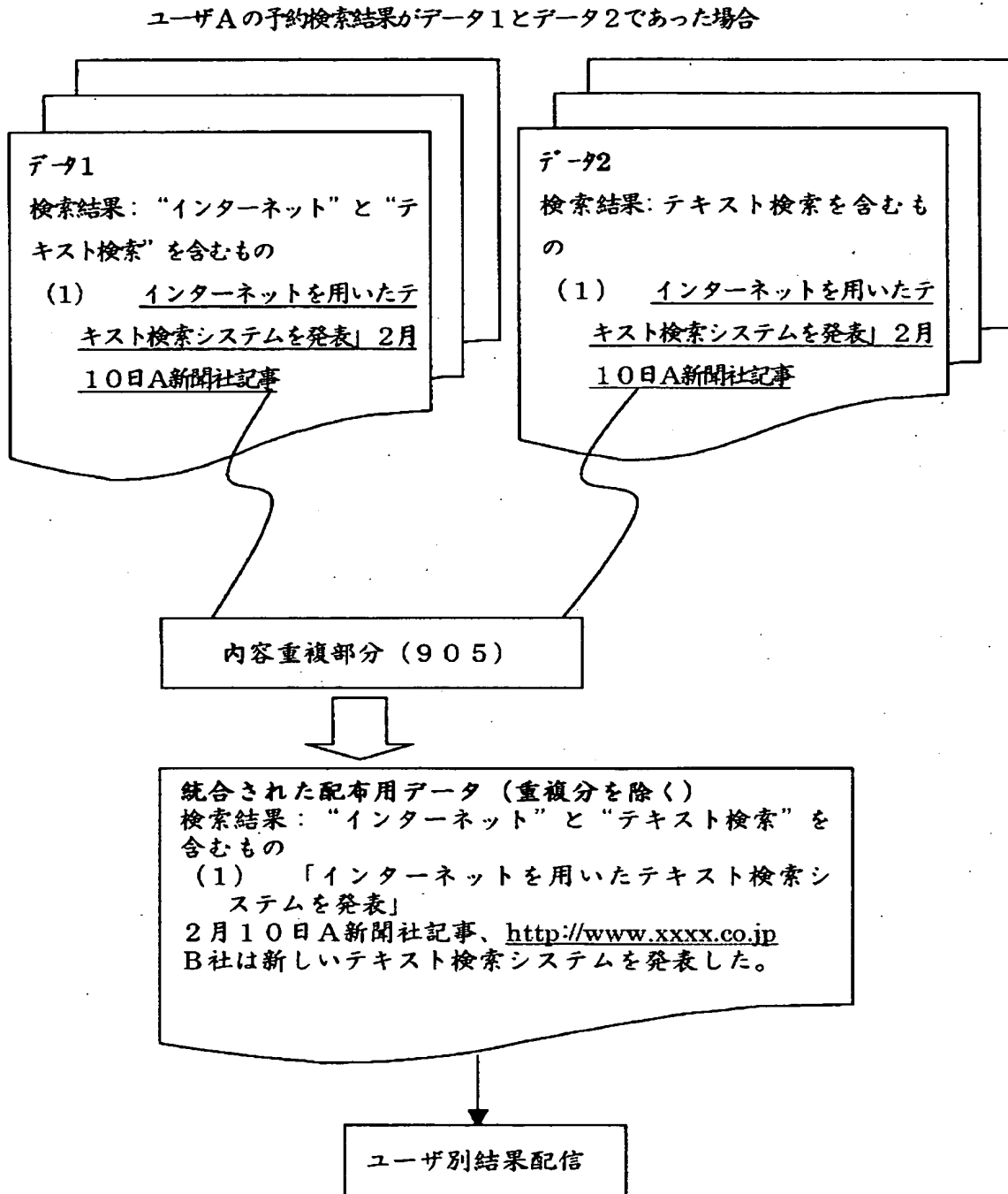
【図9】

図9 検索結果データ重複統合の仕組み



【図10】

図10 重複統合された検索結果データの内容例



【図 11】

図 11 配信用テンプレート(1101)の例

1101

検索キー:  
検索日時:  
目次: 検索条件と結果要約一覧  
検索結果:  
ヒット数:  
{番号、タイトル、作成者、作成日付、URL、概要、コメント}  
{番号、タイトル、作成者、作成日付、URL、概要、コメント}  
{番号、タイトル、作成者、作成日付、URL、概要、コメント}  
:  
終了コメント:

The diagram shows a rectangular box representing a distribution template. To the right of the box, the number '1101' is written, with a curved line pointing from it to the '検索結果:' (Search Results) section of the template. The template itself contains the following text: '検索キー:' (Search Key), '検索日時:' (Search Date/Time), '目次: 検索条件と結果要約一覧' (Table of Contents: Search Conditions and Results Summary), '検索結果:' (Search Results), 'ヒット数:' (Hit Count), a list of three items in curly braces: '{番号、タイトル、作成者、作成日付、URL、概要、コメント}' (Number, Title, Creator, Creation Date, URL, Summary, Comment), a vertical ellipsis ':', and '終了コメント:' (End Comment).

【図12】

図12 配信用テンプレートを用いた配信メールの例

ユーザA宛てのメール (1201)

1201

検索キー：インターネット、テキスト検索

検索日時：2001年3月1日

目次：(1) インターネットを用いたテキスト検索システム  
(2) インターネット上のテキスト検索の新技法

=====  
検索結果：“インターネット”と“テキスト検索”を含むもの  
ヒット数：2件  
(1) 「インターネットを用いたテキスト検索システムを発表」  
2月10日A新聞社記事、<http://www.xxxx.co.jp>  
B社は新しいテキスト検索システムを発表した。  
=====  
(2) 「インターネット上のテキスト検索の新技法」  
2月11日C学会誌、<http://www.xx.org>  
従来のテキスト検索と異なり、新しいテキスト検索システムではあ  
いまいな意味でも検索可能となる。  
=====  
検索結果：“テキスト検索”のみを含むもの  
ヒット数：1件  
(3) 「知識管理における全文テキスト検索のありかた」  
2月10日A新聞社記事、<http://www.xxxx.co.jp>  
知識管理に対する関心が高まるなか、全文検索についてD学会から  
発表がなされた。  
=====  
検索結果：“インターネット”のみを含むもの  
ヒット件数：1523件  
コメント：メール容量オーバーのため、検索結果DB格納エリア1  
を直接参照してください。  
=====  
終了メッセージ：以上が検索結果です。



【書類名】 要約書

【要約】

【課題】

文書情報の予約検索において、複数ユーザの登録した検索条件に重複があった場合にこれを排除して効率のよい検索を行うことのできる検索システムおよび検索サービスを提供する。

【解決手段】

予約検索を行うに際し、ユーザ間で検索条件の登録内容に重複があった場合には重複を統合して検索処理を実行し、得られた検索結果について、ユーザ間の重複分について各ユーザが所望する検索条件登録に合致するように情報の複製を行って各ユーザの検索条件登録ごとに展開するとともに、展開した内容中に重複があった場合はこれを排除して統合し、ユーザが所望する形態に編集して編集結果をユーザが予め登録したメールアドレスなどの配信先へ配信する。

【選択図】 図 1

認定・付加情報

特許出願の番号	特願2001-091636
受付番号	50100442677
書類名	特許願
担当官	第七担当上席 0096
作成日	平成13年 3月29日

<認定情報・付加情報>

【提出日】	平成13年 3月28日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日	1990年 8月31日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台4丁目6番地
氏 名	株式会社日立製作所